INFLECTION POINT PHASE FIELD HIGHER-ORDER ACTIVE CONTOUR MODEL FOR NETWORK EXTRACTION FROM VHR SATELLITE IMAGES

Aymen El Ghoul, Ian H. Jermyn, and Josiane Zerubia

ARIANA - Joint Team-Project INRIA/CNRS/UNSA
2004 route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex, France
email: {aymen.el_ghoul, ian.jermyn, josiane.zerubia}@sophia.inria.fr
web: www-sop.inria.fr/ariana

ABSTRACT

The segmentation of networks is important in several imaging domains, and models incorporating prior shape knowledge are often essential for the automatic performance of this task. We incorporate such knowledge via phase fields and higher-order active contours (HOACs). In this paper: we introduce an improved prior model, the phase field HOAC 'inflection point' model of a network; we present an improved data term for the segmentation of road networks; we confirm the robustness of the resulting model to choice of gradient descent initialization; and we illustrate these points via road network extraction results on VHR satellite images.

1. INTRODUCTION

Remote sensing imagery plays an important role in many areas, including cartography, and the study and maintenance of the environment. This is particularly true of very high resolution (VHR) satellite images, which offer an enormous richness of information, if only it could be extracted. In this paper, we focus on road network extraction in rural areas from VHR multi-spectral satellite images.

The core contribution of the present paper is to refine the phase field higher-order active contour (HOAC) model introduced by [6] using the stability analysis described in [3]: we insist that a branch of the road network be an inflection point of the energy rather than a local minimum. This has two effects: it reduces to one the number of free parameters in the prior model, and it eliminates 'phantom roads' in the background region.

In addition, we test two models for multi-spectral image data in both background and foreground: a multivariate Gaussian model, and a mixture of two multivariate Gaussians. While the former behaves better as a maximum likelihood estimator, the latter invariably outperforms it when coupled with the phase field HOAC prior. We also test the initialization sensitivity of the model, always an important issue when dealing with gradient descent optimization methods as we do here. We find that due to the large amount of prior information contained in the model, many potential local minima are eliminated, and the result is very insensitive to the initialization except for one extreme worst case.

1.1 Background

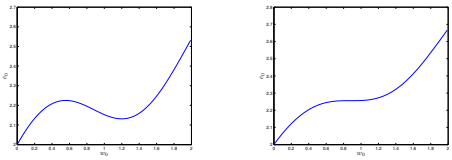
Road network extraction is a difficult problem, because such networks are frequently not distinguishable from the background by local image measurements alone. Rather, it is the 'shape' of the region R in the image domain occupied by the network that distinguishes them: networks occupy regions in the image domain composed of a set of branches that join together at junctions. We adopt a probabilistic approach and MAP estimation to formulate the problem: we estimate the region R by maximizing a posterior probability, $\mathrm{P}(R|I,K)$ where I is the original image and K represents prior knowledge about R and the relation between R and I. As usual, this can be written up to an irrelevant normalization as the product of a likelihood $\mathrm{P}(I|R,K)$ and a prior $\mathrm{P}(R|K)$. In practice, we deal with negative log probabilities: we minimize a total energy $E(R,I)$ that is the sum of a likelihood term $E_{\mathrm{I}}(I,R)$ and a prior term $E_{\mathrm{P}}(R)$.

There has been a significant amount of work on shape modelling for segmentation, but most of this work uses priors that model shape by comparing the region sought to reference regions [1, 8]. This works well for many applications, but is not appropriate when the region sought can have an arbitrary topology, as in the case of networks. To model such regions, 'higher-order active contours' (HOACs) were introduced [7]. Unlike the original active contours [5] and their successors, this new generation of active contours incorporates non-trivial shape information about a region via long-range interactions between region boundary points.

The HOAC prior energy defined in [7] was used to model network-shaped regions. However, the same energy can be used to model other families of regions too, notably a 'gas of circles' [4]), simply by varying the model parameters. Thus, in order to use the model for network segmentation, one has to know which parameter ranges lead to stable networks, as opposed, for example, to stable circles. In [3], the stability to small perturbations of a network branch, abstracted as a long bar, was analysed via a Taylor series expansion up to second order. The result is a phase diagram that shows which parameter values give rise to stable networks.

The contour representation suffers from a number of drawbacks, both for classical active contours in general, and for HOACs in particular. To overcome these drawbacks, in [6] HOACs were reformulated as nonlocal 'phase field' models. This formulation facilitates model analysis and implementation, allows a 'neutral' initialization and complete topological freedom, and results in reduced execution times, sometimes by an order of magnitude.

In section 2, we briefly recall the HOAC model for networks, and the stability analysis. We then explain, in section 2.2, the rationale and implementation of the main contribution of the present paper: the 'inflection point model'. In section 2.3, we briefly recall the phase field formulation of HOACs, and then show how to construct a phase field version of the inflection point model. In section 3, we describe the two data models we will test. In section 4, we describe



(a) (b)

Figure 1: 1(a) and 1(b) plot e_0 against bar width w_0, with $(\alpha_C, \beta_C) = (0.8, 0.53)$ and $(0.7, 0.363)$ respectively, giving a minimum at $w_0^* = 1.2$ and an inflection point at $w_0^* = 0.88$.

experiments that test the data models, confirm the theoretical rationale for and practical improvements obtained from the inflection point model, and that demonstrate the robustness of the results to the initialization. We conclude in section 5.

2. INFLECTION POINT MODELS OF NETWORKS

The prior HOAC energy introduced in [7] is

$$E_{C,\mathrm{G}}(R) = \lambda_C L(R) + \alpha_C A(R) - \frac{\beta_C}{2} \iint_{(S^1)^2} dt\, dt'\, \dot\gamma(t) \cdot \dot\gamma(t')\, \Psi\left(\frac{|\gamma(t) - \gamma(t')|}{d}\right), \quad (1)$$

where R is a region in the image domain; L is region boundary length; A is region area; γ is an embedding representing the region boundary, parameterized by $t \in S^1$; $\dot\gamma$ is its derivative, *i.e.* the tangent vector field; $\Psi(z)$ defines the interaction between two boundary points separated by a distance zd; and λ_C, α_C, β_C, and d are real parameters. The long-range interaction is responsible for the prior shape information.

This HOAC prior energy was used in [7] to model networks. However, it was soon discovered that it could be used to model other families of regions too, notably a 'gas of circles' [4], simply by varying the model parameters. Thus, in order to use the model for network segmentation, it is important to know which parameter ranges lead to stable networks.

2.1 Stable networks

In [3], a stability analysis of a network branch, abstracted as a long bar, was described. The energy $E_{C,\mathrm{G}}$ was expanded about a long bar configuration to second order in a small perturbation $\delta\gamma$ (expressed in the Fourier basis for simplicity):

$$e_{\mathrm{G}}^{(2)}(\gamma) = e_0 + e_1\,(a_{1,0} - a_{2,0}) + \frac{1}{2}\sum_k a_k^\dagger\, e_2\, a_k\,, \quad (2)$$

where $a_k = (a_{1,k}^*, a_{2,k})$ and $a_{i,k}$ is the Fourier coefficient of frequency k corresponding to the side of the bar labelled by $i \in \{1,2\}$; $\dagger$ and $*$ indicate respectively Hermitian and complex conjugates. $e_0(w_0)$ is the energy per length unit of a long bar of width w_0, while $e_1(w_0) = \partial e_0(w_0)/\partial w_0$. $e_2(w_0,k)$ is, for each frequency k, a symmetric 2×2 matrix whose diagonal and off-diagonal terms, e_{20} and e_{21}, express the self-energy of perturbations of one side, and the interaction between perturbations of opposite sides of the bar respectively. For future use, we define dimensionless parameters $\hat w_0 = w_0/d$, $\hat\alpha = \alpha_C d/\lambda_C$ and $\hat\beta = \beta_C d/\lambda_C$, and without loss

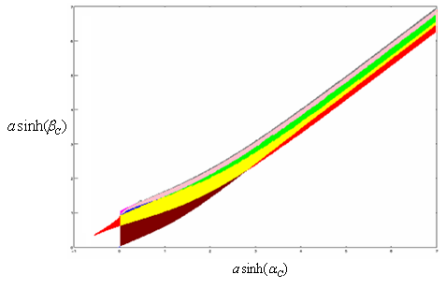


Figure 2: Phase diagram. Maroon, red, yellow, green, white, blue, pink, grey, magenta correspond respectively to B+, C+, B+ C+, B+ C-, UB UC, B- C+, B- C-, C- and B-; B, C, U, + and - refer respectively to bar, circle, unstable, positive energy and negative energy.

of generality, we put $\lambda_C = d = 1$. Thus the problem is reduced to finding the ranges of α_C and β_C allowing a stable bar. The bar is stable (*i.e.* an energy minimum) if and only if: $e_1(\alpha_C, w_0, \beta_C) = 0$, which determines β_C in terms of α_C and w_0; and $e_2(\alpha_C, w_0, \beta_C, k)$ is positive definite for all k, *i.e.* its eigenvalues $\lambda_{\pm,k} = e_{20}(k) \pm e_{21}(k)$, are positive for all k. These conditions give upper and lower bounds on the parameter values. The result of the analysis, when combined with a similar analysis for the 'gas of circles' model, is illustrated in the phase diagram shown in figure 2. The phase diagram enables parameter values to be chosen to model a particular situation. The sign of e_0 is important: if $e_0 < 0$ then the bar lengthens to minimize the total energy, while if $e_0 > 0$ then the bar shrinks until it disappears. The first situation is undesirable because gradient descent tends to create arbitrary network branches to minimize the total energy. Suitable parameter values therefore lie in the maroon region of the phase diagram, which gives a stable bar with positive energy per unit length and no stable circles.

2.2 HOAC inflection point network model

Figure 1(a) shows the energy per unit length $e_0(w_0)$ plotted against bar width w_0 for a particular parameter setting from the maroon region with a stable width $\hat w_0^* = 1.2$. Parameter settings that produce energy curves like figure 1(a) have a disadvantage when minimized using gradient descent. Imagine an area of background in the image, and a network branch formed there by the vagaries of gradient descent. Because it lies in the background, and assuming the data model is reasonable, there will be a force inwards on the branch, tending to make it shrink and disappear. This is as it should be. However, if the width of the branch lies in the basin of attraction of the stable width, there is a threshold that the force has to surmount if it is to push the branch over the maximum shown in figure 1(a), and down to zero width. the result is the formation of 'phantom roads', false positives that cannot disappear due to the stability of the network branch.

Global optimization algorithms are one way to avoid local minima, but unfortunately our problem is NP-hard. We choose a slightly less ambitious approach: we change the energy functional to avoid the creation of these local minima while preserving as much prior knowledge as possible. This problem can be solved by constraining the parameters so that the energy function has an inflection point at a desired width

w_0 (*i.e.* $\lambda_-(w_0,0)=0$) rather than a minimum. Figure 1(b) shows a plot of energy per unit length versus width for a parameter setting that gives an inflection point. Such inflection points lie on the lower edge of the coloured area of the phase diagram, with $\hat{\alpha}$ values in the range $[0,0.9083]$. The value of $\hat{w}_0 = 0.88$ all along this line, and this is the only value that allows an inflection point. We test this potential solution to the 'phantom road' problem in section 4.

2.3 Phase fields

A phase field ϕ is a real-valued function on the image domain Ω. A phase field determines a region by the map $\zeta_z(\phi) = \{x \in \Omega : \phi(x) > z\}$ where z is a given threshold. The basic phase field energy is [6]

$$E_0(\phi) = \int_\Omega dx \left\{ \frac{D}{2} \partial\phi \cdot \partial\phi + \lambda\left(\frac{\phi^4}{4} - \frac{\phi^2}{2}\right) + \alpha\left(\phi - \frac{\phi^3}{3}\right) \right\}. \quad (3)$$

The ultralocal terms define the stable phases: they have two minima, at -1 and $+1$, which correspond to the background and the network respectively. The derivative term ensures the smoothness of ϕ, giving it an interface of finite width, w, around the boundary ∂R. To introduce prior shape information, a nonlocal term is then added, as with HOACs, to give a total energy $E_G = E_0 + E_{NL}$, where

$$E_{NL}(\phi) = -\frac{\beta}{2} \iint_{\Omega^2} dx\, dx'\, \partial\phi(x) \cdot \partial\phi(x')\, \Psi\left(\frac{|x-x'|}{d}\right). \quad (4)$$

Let $\phi_R = \arg\min_{\phi:\, \zeta_z(\phi)=R} E_G(\phi)$. Then it can be shown that $E_G(\phi_R) \simeq E_{C,G}(R)$, thus enabling the use of phase fields instead of HOACs. The parameters of the phase field model can be expressed in terms of those of the contour model: $\alpha = 0.75\alpha_C$, $\beta = 0.25\beta_C$, $D = 0.25 w \lambda_C$ and $\lambda = 15\lambda_C(1 + \sqrt{1 - 4w^2(\alpha_C/\lambda_C)^2/5})/(8w)$.

2.4 Phase field inflection point network model

To model networks with the phase field model, we first select parameter values for the contour using the phase diagram. In practice, this means fixing w_0 (which is an application-determined physical parameter), and then selecting values of $\hat{\alpha}$ and $\hat{\beta}$ from the maroon region of the phase diagram. These give $\hat{w}_0$, which gives the required d, and hence α_C/λ_C, which is upper-bounded so that λ is real. A choice of λ_C then gives the actual values of the parameters in $E_{C,G}$. These are then converted using the equations at the end of the last section (we choose $w = 3$).

Once we impose the inflection point condition, $\hat{w}_0$ is fixed, and $\hat{\alpha}$ is sufficient to determine $\hat{\beta}$ and hence all the other parameters except λ_C. However, the inflection point condition constrains the parameters to a codimension 1 set in parameter space, so that a generic change in the parameters, however small, will violate the condition. One can then wonder how well this condition is preserved when the parameters are converted from contour to phase field, especially since this conversion is based on a relatively crude approximation to ϕ_R. In practice, numerical experiments show that the inflection point condition is well preserved, with configurations at the inflection point remaining stationary to subpixel accuracy over thousands of iterations of gradient descent.

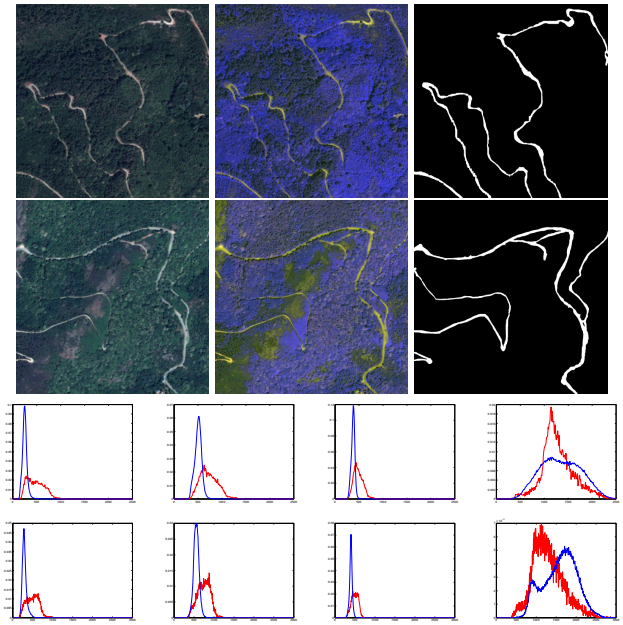


Figure 3: Top row, from left to right: the R-G-B bands of a multispectral satellite image; the G-R-IR bands of the same image; the corresponding manually extracted road network mask. Second row: similar, for a second image. Third row, from, left to right: histograms of the network (red) and background (blue) regions of the R, G, B, and IR bands of the image in the top row; Fourth row: similar, for second image. (Images ©DigitalGlobe, CNES processing, images acquired via ORFEO Accompaniment Program)

3. LIKELIHOOD ENERGY AND ENERGY MINIMIZATION

So far we have spoken only of the prior energy, $E_P = E_G$. In this section, we focus on the likelihood energy E_I. Figure 3 shows two, 120cm resolution, multi-spectral test images (red, green, blue and infrared bands), together with manually extracted road network masks, and the histograms of the network and the background for each band. We now define two data models: MG and MMG.

3.1 Data energy term

We assume that $P(I|R,K) = P(I_R|R,\theta_R,K)P(I_{\bar{R}}|\bar{R},\theta_{\bar{R}},K)$, where I_R and $I_{\bar{R}}$ are the restrictions of the image to the network R and the background $\bar{R}$ respectively, and θ_R and $\theta_{\bar{R}}$ are the corresponding model parameters (which previously were included in the generic K). We further assume that the image values at different pixels are independent given these parameters. Taking negative logarithms, and using $\phi_\pm = (1 \pm \phi)/2$ to restrict integrations to the network or background respectively, gives the following likelihood energy:

$$\begin{aligned} E_I(\phi) &= -\int_\Omega dx \left\{ \ln(\pi_{\theta_R}(I(x)))\, \phi_+(x) + \ln(\pi_{\theta_{\bar{R}}}(I(x)))\, \phi_-(x) \right\} \\ &= -\int_\Omega dx\, \frac{\phi(x)}{2} \left\{ \ln(\pi_{\theta_R}(I(x))) - \ln(\pi_{\theta_{\bar{R}}}(I(x))) \right\} + k\,, \end{aligned}$$

where k is a ϕ-independent constant, which we drop.

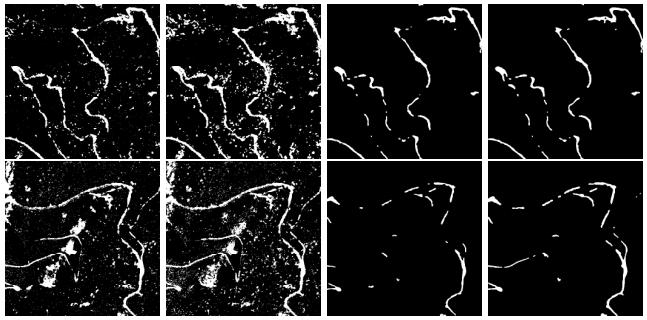


Figure 4: Classification results for the two images in figure 3, from left to right in each row: ML using MG, ML using MMG, MAP using MG, and MAP using MMG. From 3rd to 4th column and from top to bottom: $(w_0, \hat{\alpha}, \lambda_C) = (2, 0.7646, 15)$, $(2, 0.8385, 5)$, $(2, 0.8385, 20)$, and $(2, 0.6169, 10)$.

3.2 Multivariate Gaussian model

After simple calculations, the MG model can be written as

$$E_{\mathrm{I}}(\phi) = \frac{1}{4}\int_{\Omega} dx \left\{ (I(x)-\mu)^T \Sigma^{-1} (I(x)-\mu) - (I(x)-\bar{\mu})^T \bar{\Sigma}^{-1} (I(x)-\bar{\mu}) + \ln \frac{|\Sigma|}{|\bar{\Sigma}|} \right\} \phi(x) \, ,$$

where μ and $\bar{\mu}$ are the mean vectors of the bands of I_R and $I_{\bar{R}}$ respectively; Σ and $\bar{\Sigma}$ are the covariance matrices of I_R and $I_{\bar{R}}$ respectively. $\theta = (\mu, \bar{\mu}, \Sigma, \bar{\Sigma})$ is learnt from the original images and their masks using maximum likelihood.

3.3 Multivariate mixture of two Gaussian model

The MMG model is designed to take into account the heterogeneity in the appearance of the network produced by occlusions. It takes the form

$$E_{\mathrm{I}}(\phi) = -\frac{1}{2}\int_{\Omega} dx \left\{ \ln \sum_{i=1}^{2} p_i |2\pi\Sigma_i|^{-1/2} e^{-\frac{1}{2}(I(x)-\mu_i)^T \Sigma_i^{-1} (I(x)-\mu_i))} - \ln \sum_{i=1}^{2} \bar{p}_i |2\pi\bar{\Sigma}_i|^{-1/2} e^{-\frac{1}{2}(I(x)-\bar{\mu}_i)^T \bar{\Sigma}_i^{-1} (I(x)-\bar{\mu}_i))} \right\} \phi(x) \, .$$

where p_i and $\bar{p}_i$ weight the two Gaussian components for I_R and $I_{\bar{R}}$ respectively. $\theta = (p_1, p_2, \bar{p}_1, \bar{p}_2, \mu_1, \mu_2, \bar{\mu}_1, \bar{\mu}_2, \Sigma_1, \Sigma_2, \bar{\Sigma}_1, \bar{\Sigma}_2)$ is learnt from the original images and their masks using maximum likelihood combined with the EM algorithm [2].

4. EXPERIMENTS AND DISCUSSION

The total phase field energy to minimize is $E(\phi) = E_{\mathrm{G}}(\phi) + E_{\mathrm{I}}(\phi)$. We use gradient descent to seek minima [6].

4.1 MG model vs. MMG model

We compare the segmentation performance of the MG and MMG data models. We begin by looking at the performance of the two models using maximum likelihood classification, *i.e.* with E_{G} set to zero. Results on the two images in figure 3 are shown in the first two columns of figure 4. In this case, the MMG model performs worse than the MG model, since it allows some parts of the background to be classified as network. With the addition of the prior, however, the results using the MMG model are better than those of the MG model on both images, as shown in the last two columns of figure 4. See table 1 for quantitative evaluations.

		Completeness = TP / (TP+FN)	Correctness = TP / (TP+FP)	Quality = TP / (TP+FP+FN)
ML	MG	0.7343	0.4923	**0.4179**
	MGM	0.8152	0.3467	0.3214
MAP	MG	0.5962	0.7955	0.5170
	MGM	0.5982	0.8028	**0.5216**
ML	MG	0.6839	0.3754	**0.3199**
	MGM	0.7370	0.3338	0.2983
MAP	MG	0.3275	0.9405	0.3208
	MGM	0.4730	0.9282	**0.4563**
EMM	MMG	0.7591	0.5798	0.4897
EIPM	MMG	0.5982	0.8028	**0.5216**

Table 1: Quantitative evaluations of the experiments. T, F, P, and N correspond to true, false, positive, and negative respectively.

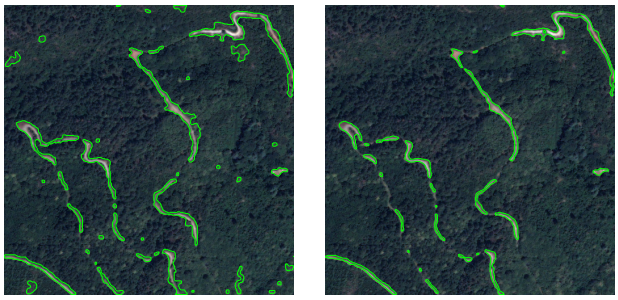


Figure 5: Left: segmentation result using parameter values selected from the maroon zone, $(w_0, \hat{\alpha}, \lambda_C) = (4, 0.2013, 5)$. Right: segmentation result using parameter values leading to an inflection point at the desired bar width, $(w_0, \hat{\alpha}, \lambda_C) = (2, 0.7646, 15)$.

4.2 Inflection point model

We compare the segmentation performance of the previous, energy minimum model (EMM) and the new energy inflection point model (EIPM). The first is obtained using parameter values from the maroon region of the phase diagram given by figure 2, and gives an energy per unit length as a function of width as shown in figure 1(a), while the second uses the parameter constraints detailed in section 2.4 to create an inflection point model, with an energy per unit length as a function of width as shown in figure 1(b). The idea is to avoid false positives in the background by rendering a network configuration only marginally stable in the absence of supporting image data. The results are shown in figure 5: the false positives are indeed eliminated without creating false negatives. See table 1 for quantitative evaluations.

4.3 Robustness of the algorithm to initial conditions

Initialization dependence of the final result, *i.e.* becoming trapped in a local minimum, is a drawback of deterministic descent algorithms. However, it might be hoped that with sufficient prior knowledge built into the model, the entropy of the probability distribution would be reduced enough to eliminate most, if not all local minima, and thereby reduce or remove initialization dependence. To test this, we examine the convergence of the algorithm using different initializations for the phase field ϕ:

- to the constant value $\phi_0 = \alpha/\lambda$, which corresponds to the maximum of V, which is the threshold z, and hence to all of Ω being boundary (the neutral initialization, NI);
- to the constant value -1, corresponding to all of Ω being background;
- to the constant value $+1$, corresponding to all of Ω being foreground;
- when values of ϕ are sampled independently from a uniform distribution on $[-1,1]$ (UR);
- to the ML result;
- to $1-$ the ML result;
- to the ML result scaled linearly towards ϕ_0.

Figure 6 shows segmentations of the two images in figure 3, each with two different parameter settings, and using the above initializations. The converged solutions agreed to within 0.5% pixel difference, except for the initial condition $+1$ for one of the two parameter settings, where the algorithm does not converge to the correct solution.

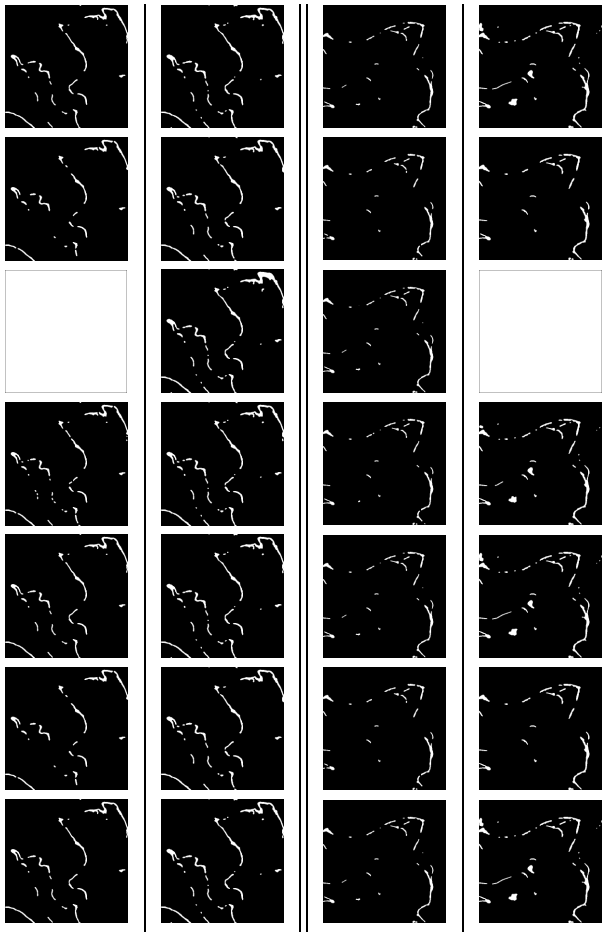


Figure 6: Segmentations of the two images in figure 3 using different initializations. From top to bottom: NI, -1, +1, UR, ML, -ML, Scaled ML. 1st image: $(w_0, \hat{\alpha}, \lambda_C) = (3, 1.2578, 30)$ for the 1st column and $(3, 1.2578, 20)$ for the 2nd column. Second image: $(2, 0.5924, 15)$ for the 3^{rd} column and $(2, 0.8385, 20)$ for the 4^{th} column. The MG data model was used with the inflection point prior model.

5. CONCLUSION

We have described a phase field HOAC model for road network extraction from VHR satellite images of rural areas. The contributions of this model are: the use of inflection point parameter values, which we show both reduces the number of free prior parameters and eliminates false positives; the use of multivariate mixture of two Gaussians models for foreground and background, which we show outperforms the maximum-likelihood-preferred multivariate Gaussian models when coupled with our prior model; and initialization independence despite the use of deterministic gradient descent, as shown by experiments.

Acknowledgments

The authors thank the French Space Agency (CNES) for the satellite images, and CNES, the PACA Region, the MUSCLE Network of Excellence (FP6-507752, www.muscle-noe.org), and INRIA Associated Team SHAPES (www-sop.inria.fr/ariana/Projets/Shapes/index.html) for partial financial support.

REFERENCES

[1] D. Cremers, M. Rousson, and R. Deriche. A review of statistical approaches to level set segmentation: Integrating color, texture, motion and shape. *IJCV*, 72(2):195–215, 2007.

[2] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):1–38, 1977.

[3] A. El Ghoul, I. H. Jermyn, and J. Zerubia. Phase diagram of a long bar under a higher-order active contour energy: application to hydrographic network extraction from VHR satellite images. In *Proc. ICPR*, Tampa, Florida, December 2008.

[4] P. Horvath, I.H. Jermyn, Z. Kato, and J. Zerubia. A higher-order active contour model for tree detection. In *Proc. ICPR*, Hong Kong, August 2006.

[5] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *IJCV*, 1(4):321–331, 1988.

[6] M. Rochery, I. H. Jermyn, and J. Zerubia. Phase field models and higher-order active contours. In *Proc. ICCV*, Beijing, China, October 2005.

[7] M. Rochery, I.H. Jermyn, and J. Zerubia. Higher order active contours. *IJCV*, 69(1):27–42, 2006.

[8] A. Srivastava, S. Joshi, W. Mio, and X. Liu. Statistical shape analysis: Clustering, learning, and testing. *Trans. PAMI*, 27(4):590–602, 2003.